

#6

OIPE

## RAW SEQUENCE LISTING

DATE: 06/25/2001

PATENT APPLICATION: US/09/777,566A

TIME: 11:28:44

Input Set : A:\DIVER1370-6.ST25.txt

Output Set: N:\CRF3\06252001\I777566A.raw

ENTERED

3 <110> APPLICANT: DIVERSA CORPORATION  
4 SHORT, Jay  
5 KRETZ, Keith  
7 <120> TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF  
9 <130> FILE REFERENCE: DIVER1370-6  
11 <140> CURRENT APPLICATION NUMBER: US 09/777,566A  
C--> 12 <141> CURRENT FILING DATE: 2001-06-11  
14 <150> PRIOR APPLICATION NUMBER: US 09/318,528  
15 <151> PRIOR FILING DATE: 1999-05-25  
17 <150> PRIOR APPLICATION NUMBER: US 09/291,931  
18 <151> PRIOR FILING DATE: 1999-04-13  
20 <150> PRIOR APPLICATION NUMBER: US 09/259,214  
21 <151> PRIOR FILING DATE: 1999-03-01  
23 <150> PRIOR APPLICATION NUMBER: US 08/910,798  
24 <151> PRIOR FILING DATE: 1997-08-13  
26 <160> NUMBER OF SEQ ID NOS: 4  
28 <170> SOFTWARE: PatentIn version 3.0  
30 <210> SEQ ID NO: 1  
31 <211> LENGTH: 1323  
32 <212> TYPE: DNA  
33 <213> ORGANISM: Escherichia coli  
35 <220> FEATURE:  
36 <221> NAME/KEY: CDS  
37 <222> LOCATION: (1)..(1320)  
39 <220> FEATURE:  
40 <221> NAME/KEY: misc\_feature  
41 <222> LOCATION: (1)..(1323)  
42 <223> OTHER INFORMATION: n is any nucleotide  
45 <400> SEQUENCE: 1  
46 atg aaa gcg atc tta atc cca ttt tta tct ctt ctg att ccg tta acc 48  
47 Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr  
48 1 5 10 15  
50 ccg caa tct gca ttc gct cag agt gag ccg gag ctg aag ctg gaa agt 96  
51 Pro Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser  
52 20 25 30  
54 gtg gtg att gtc agt cgt cat ggt gtg cgt gct cca acc aag gcc acg 144  
55 Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr  
56 35 40 45  
58 caa ctg atg cag gat gtc acc cca gac gca tgg cca acc tgg ccg gta 192  
59 Gln Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val  
60 50 55 60  
W--> 62 aaa ctg ggt tgg ctg aca ccg cgn ggt ggt gag cta atc gcc tat ctc 240  
63 Lys Leu Gly Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu  
64 65 70 75 80  
66 gga cat tac caa cgc cag cgt ctg gta gcc gac gga ttg ctg gcg aaa 288  
67 Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys  
68 85 90 95

## RAW SEQUENCE LISTING

DATE: 06/25/2001

PATENT APPLICATION: US/09/777,566A

TIME: 11:28:44

Input Set : A:\DIVER1370-6.ST25.txt

Output Set: N:\CRF3\06252001\I777566A.raw

70	aag ggc tgc ccg cag tct ggt cag gtc gcg att att gct gat gtc gac	336
71	Lys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp	
72	100 105 110	
74	gag cgt acc cgt aaa aca ggc gaa gcc ttc gcc gcc ggg ctg gca cct	384
75	Glu Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro	
76	115 120 125	
78	gac tgt gca ata acc gta cat acc cag gca gat acg tcc agt ccc gat	432
79	Asp Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp	
80	130 135 140	
82	ccg tta ttt aat cct cta aaa act ggc gtt tgc caa ctg gat aac gcg	480
83	Pro Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala	
84	145 150 155 160	
86	aac gtg act gac gcg atc ctc agc agg gca gga ggg tca att gct gac	528
87	Asn Val Thr Asp Ala Ile Leu Ser Arg Ala Gly Gly Ser Ile Ala Asp	
88	165 170 175	
90	ttt acc ggg cat cgg caa acg gcg ttt cgc gaa ctg gaa cgg gtg ctt	576
91	Phe Thr Gly His Arg Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu	
92	180 185 190	
94	aat ttt ccg caa tca aac ttg tgc ctt aaa cgt gag aaa cag gac gaa	624
95	Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu	
96	195 200 205	
98	agc tgt tca tta acg cag gca tta cca tcg gaa ctc aag gtg agc gcc	672
99	Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala	
100	210 215 220	
102	gac aat gtc tca tta acc ggt gcg gta agc ctc gca tca atg ctg acg	720
103	Asp Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr	
104	225 230 235 240	
106	gag ata ttt ctc ctg caa caa gca cag gga atg ccg gag ccg ggg tgg	768
107	Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp	
108	245 250 255	
110	gga agg atc acc gat tca cac cag tgg aac acc ttg cta agt ttg cat	816
111	Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His	
112	260 265 270	
114	aac gcg caa ttt tat ttg cta caa cgc acg cca gag gtt gcc cgc agc	864
115	Asn Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser	
116	275 280 285	
118	cgc gcc acc ccg tta ttg gat ttg atc atg gca gcg ttg acg ccc cat	912
119	Arg Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His	
120	290 295 300	
122	cca ccg caa aaa cag gcg tat ggt gtg aca tta ccc act tca gta ctg	960
123	Pro Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu	
124	305 310 315 320	
126	ttt att gcc gga cac gat act aat ctg gca aat ctc ggc ggc gca ctg	1008
127	Phe Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu	
128	325 330 335	
130	gag ctc aac tgg acg ctt ccc ggt cag ccg gat aac acg ccg cca ggt	1056
131	Glu Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly	
132	340 345 350	
134	ggt gaa ctg gtg ttt gaa cgc tgg cgt ccg cta agc gat aac agc cag	1104

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/777,566A

DATE: 06/25/2001

TIME: 11:28:44

Input Set : A:\DIVER1370-6.ST25.txt

Output Set: N:\CRF3\06252001\I777566A.raw

```

135 Gly Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln
136          355          360          365
138 tgg att cag gtt tcg ctg gtc ttc cag act tta cag cag atg cgt gat      1152
139 Trp Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp
140      370          375          380
142 aaa acg ccg ctg tca tta aat acg ccg ccc gga gag gtg aaa ctg acc      1200
143 Lys Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr
144 385          390          395          400
146 ctg gca gga tgt gaa gag cga aat gcg cag ggc atg tgt tcg ttg gca      1248
147 Leu Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala
148          405          410          415
150 ggt ttt acg caa atc gtg aat gaa gca cgc ata ccg gcg tgc agt ttg      1296
151 Gly Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu
152          420          425          430
154 aga tct cat cac cat cac cat cac taa      1323
155 Arg Ser His His His His His His
156          435          440
159 <210> SEQ ID NO: 2
160 <211> LENGTH: 440
161 <212> TYPE: PRT
162 <213> ORGANISM: Escherichia coli
164 <220> FEATURE:
165 <221> NAME/KEY: misc_feature
166 <222> LOCATION: (1)..(1323)
167 <223> OTHER INFORMATION: n is any nucleotide
169 <400> SEQUENCE: 2
171 Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr
172 1          5          10          15
175 Pro Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser
176          20          25          30
179 Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr
180          35          40          45
183 Gln Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val
184          50          55          60
187 Lys Leu Gly Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
188 65          70          75          80
191 Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys
192          85          90          95
195 Lys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp
196          100          105          110
199 Glu Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro
200          115          120          125
203 Asp Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp
204          130          135          140
207 Pro Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala
208 145          150          155          160
211 Asn Val Thr Asp Ala Ile Leu Ser Arg Ala Gly Gly Ser Ile Ala Asp
212          165          170          175
215 Phe Thr Gly His Arg Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu

```

## RAW SEQUENCE LISTING

DATE: 06/25/2001

PATENT APPLICATION: US/09/777,566A

TIME: 11:28:44

Input Set : A:\DIVER1370-6.ST25.txt

Output Set: N:\CRF3\06252001\I777566A.raw

```

216          180          185          190
219 Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu
220          195          200          205
223 Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala
224          210          215          220
227 Asp Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr
228 225          230          235          240
231 Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp
232          245          250          255
235 Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His
236          260          265          270
239 Asn Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser
240          275          280          285
243 Arg Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His
244          290          295          300
247 Pro Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu
248 305          310          315          320
251 Phe Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu
252          325          330          335
255 Glu Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly
256          340          345          350
259 Gly Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln
260          355          360          365
263 Trp Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp
264          370          375          380
267 Lys Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr
268 385          390          395          400
271 Leu Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala
272          405          410          415
275 Gly Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu
276          420          425          430
279 Arg Ser His His His His His His
280          435          440

```

283 &lt;210&gt; SEQ ID NO: 3

284 &lt;211&gt; LENGTH: 49

285 &lt;212&gt; TYPE: DNA

286 &lt;213&gt; ORGANISM: Artificial sequence

288 &lt;220&gt; FEATURE:

289 &lt;223&gt; OTHER INFORMATION: Primer for PCR

291 &lt;400&gt; SEQUENCE: 3

292 gtttctgaat tcaaggagga atttaaataa aagcgatctt aatcccatt

49

295 &lt;210&gt; SEQ ID NO: 4

296 &lt;211&gt; LENGTH: 33

297 &lt;212&gt; TYPE: DNA

298 &lt;213&gt; ORGANISM: Artificial sequence

300 &lt;220&gt; FEATURE:

301 &lt;223&gt; OTHER INFORMATION: Primer for PCR

303 &lt;400&gt; SEQUENCE: 4

304 gtttctggat ccttacaac tgcacgccg tat

33

VERIFICATION SUMMARY

DATE: 06/25/2001

PATENT APPLICATION: US/09/777,566A

TIME: 11:28:45

Input Set : A:\DIVER1370-6.ST25.txt

Output Set: N:\CRF3\06252001\I777566A.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:62 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1